

PRODUCT OF RANDOM STOCHASTIC MATRICES*

BY BEHROUZ TOURI AND ANGELIA NEDIĆ

University of Illinois at Urbana-Champaign

It is well-known that for any aperiodic and irreducible stochastic matrix A , $\lim_{k \rightarrow \infty} A^k$ exists and it is a rank one stochastic matrix. We show that a generalization of this result not only holds for inhomogeneous chains of stochastic matrices but also for independent random chains of such matrices.

1. Introduction. Let $(\Omega, \mathcal{F}, \Pr)$ be a probability space and let $\{W(k)\}$ be a chain of $m \times m$ random row stochastic matrices, i.e. for all $k \geq 1$, the matrix $W(k)$ is a row stochastic almost surely and $W_{ij}(k) : \Omega \rightarrow \mathbb{R}$ is a Borel-measurable function for all $i, j \in [m]$, where $[m] = \{1, \dots, m\}$. Throughout this paper, we denote random chains by last alphabet letters such as $\{W(k)\}$ and $\{U(k)\}$, and we use the first alphabet letters such as $\{A(k)\}$ and $\{B(k)\}$ to denote deterministic stochastic chains. We will exclusively deal with row stochastic chains, so we will refer to them as stochastic chains, or just simply as chains.

Let $\{W(k)\}$ be an independent random chain. Then, we say that $\{W(k)\}$ is *strongly aperiodic* if there exists a $\gamma \in (0, 1]$ such that

$$\mathbb{E}[W_{ii}(k)W_{ij}(k)] \geq \gamma \mathbb{E}[W_{ij}(k)] \quad \text{for all } i \neq j \in [m] \text{ and all } k \geq 1.$$

Note that if $W_{ii}(k) \geq \gamma$ almost surely for all $i \in [m]$ and all $k \geq 1$, then such a chain is strongly aperiodic. Also, note that by summing both sides of the above inequality over $j \neq i$, we obtain

$$\mathbb{E}[W_{ii}(k)] \geq \mathbb{E}[W_{ii}(k)(1 - W_{ii}(k))] \geq \gamma(1 - \mathbb{E}[W_{ii}(k)]).$$

Hence, $\mathbb{E}[W_{ii}(k)] \geq \frac{\gamma}{1-\gamma}$ for all $i \in [m]$ and all $k \geq 1$. Thus, for a strongly aperiodic chain $\{W(k)\}$, the expected chain $\{\mathbb{E}[W(k)]\}$ is strongly aperiodic. It follows that a deterministic chain $\{A(k)\}$ is strongly aperiodic if and only if $A_{ii}(k) \geq \tilde{\gamma}$ for some $\tilde{\gamma} > 0$, and for all $i \in [m]$ and $k \geq 1$.

For the subsequent use, for an $m \times m$ random (or deterministic) matrix W and a non-trivial index set $S \subset [m]$ (i.e. $S \neq \emptyset$ and $S \neq [m]$), we define the quantity $W_{S\bar{S}} = \sum_{i \in S, j \in \bar{S}} W_{ij}$, where \bar{S} is the complement of the index set S .

We say that an independent random chain $\{W(k)\}$ is *balanced* if there exists some $\alpha > 0$ such that

$$(1) \quad \mathbb{E}[W_{S\bar{S}}(k)] \geq \alpha \mathbb{E}[W_{\bar{S}S}(k)] \quad \text{for all nontrivial } S \subset [m] \text{ and all } k \geq 1.$$

From this definition, it can be seen that $\alpha \leq 1$.

Finally, with a given random chain $\{W(k)\}$, let us associate a random graph $G^\infty = ([m], \mathcal{E}^\infty)$ with the vertex set $[m]$ and the edge set \mathcal{E}^∞ given by

$$\mathcal{E}^\infty(\omega) = \left\{ \{i, j\} \mid \sum_{k=1}^{\infty} (W_{ij}(k, \omega) + W_{ji}(k, \omega)) = \infty \right\}.$$

*This research was supported by the National Science Foundation under CAREER grant CMMI 07-42538.

AMS 2000 subject classifications: Primary 60F99, 60B20

We refer to G^∞ as *the infinite flow graph* of $\{W(k)\}$. By the Kolmogorov's 0-1 law, the infinite flow graph of an independent random chain $\{W(k)\}$ is almost surely equal to a deterministic graph. It has been shown that this deterministic graph is equal to the infinite flow graph of the expected chain $\{E[W(k)]\}$ ([8], Theorem 5).

For a matrix W , let W_i and W^j denote the i th row vector and the j th column vector of W , respectively. Also, for a chain $\{W(k)\}$, let $W(k : t_0) = W(k) \cdots W(t_0 + 1)$, where $k > t_0 \geq 0$ and let $W(k : k) = I$ for all $k \geq 0$. With these preliminary definitions and notation in place, we can state the main result of the current study.

THEOREM 1. *Let $\{W(k)\}$ be an independent random stochastic chain which is balanced and strongly aperiodic. Then, for any $t_0 \geq 0$ the product $W(k : t_0) = W(k) \cdots W(t_0 + 1)$ converges to a random stochastic matrix $W(\infty : t_0)$, almost surely. Furthermore, for all i, j in the same connected component of the infinite flow graph of $\{W(k)\}$, we have $W_i(\infty : t_0) = W_j(\infty : t_0)$ almost surely.*

As an immediate consequence of this result, it follows that $W(\infty : t_0)$ has rank at most τ where τ is the number of the connected components of the infinite flow graph G^∞ of $\{W(k)\}$. Thus, if G^∞ is a connected graph, the limiting random matrix $W(\infty : t_0) = \lim_{k \rightarrow \infty} W(k : t_0)$ is a rank-one random stochastic matrix almost surely, i.e. $W(\infty : t_0) = ev^T(t_0)$ almost surely for some stochastic vector $v(t_0)$. This and Theorem 1 imply that: if an independent random chain $\{W(k)\}$ is balanced and strongly aperiodic, then $\{W(k)\}$ is almost surely strongly ergodic (as defined in [3]) if and only if the infinite flow graph of $\{W(k)\}$ is connected.

2. The Dynamics System Perspective. In order to prove Theorem 1, we establish some intermediate results, some of which are applicable to a more general category of random stochastic chains, namely adapted random chains. For this, let $\{W(k)\}$ be a random chain adapted to a filtration $\{\mathcal{F}_k\}$. For an integer $t_0 \geq 0$ and a vector $v \in \mathbb{R}^m$, consider the trivial random vector $x(t_0) : \Omega \rightarrow \mathbb{R}^m$ defined by $x(t_0, \omega) = v$ for all $\omega \in \Omega$. Now, recursively define:

$$(2) \quad x(k+1) = W(k+1)x(k) \quad \text{for all } k \geq t_0.$$

Note that $x(t_0)$ is measurable with respect to the trivial σ -algebra $\{\emptyset, \Omega\}$ and hence, it is measurable with respect to \mathcal{F}_{t_0} . Also, since $\{W(k)\}$ is adapted to $\{\mathcal{F}_k\}$, it follows that for any $k > t_0$, $x(k)$ is measurable with respect to \mathcal{F}_k . We refer to $\{x(k)\}$ as a random dynamics driven by $\{W(k)\}$ started at the initial point $(t_0, v) \in \mathbb{Z}^+ \times \mathbb{R}^m$. We say that a given property holds for any dynamics $\{x(k)\}$ driven by $\{W(k)\}$ if that property holds for any initial point $(t_0, v) \in \mathbb{Z}^+ \times \mathbb{R}^m$.

Note that if $\lim_{k \rightarrow \infty} W(k : t_0) = W(\infty : t_0)$ exists almost surely, then for any initial point $(t_0, v) \in \mathbb{Z}^+ \times \mathbb{R}^m$, the random dynamics $\{x(k)\}$ converges to $W(\infty : t_0)v$ almost surely. Also, note that for any $i, j \in [m]$, we have $\lim_{k \rightarrow \infty} \|W_i(k, t_0) - W_j(k, t_0)\| = 0$ almost surely if and only if $\lim_{k \rightarrow \infty} (x_i(k) - x_j(k)) = 0$ almost surely for any initial point. When verifying the latter relation, due to linearity of the dynamics, it suffices to check that $\lim_{k \rightarrow \infty} (x_i(k) - x_j(k)) = 0$ for all the initial points of the form (t_0, e_ℓ) with $\ell \in [m]$, where $\{e_1, \dots, e_m\}$ is the standard basis for \mathbb{R}^m .

In order to study the limiting behavior of the products $W(k : t_0)$, we study the limiting behavior of the dynamics $\{x(k)\}$ driven by $\{W(k)\}$. This enables us to use the dynamics system's tools and its stability theory to draw conclusions about the limiting behavior of the matrix products $W(k : t_0)$.

2.1. Why the Infinite Flow Graph. In this section we provide a result showing the relevance of the infinite flow graph to the study of the product of stochastic matrices. Let us consider a deterministic chain $\{A(k)\}$ of stochastic matrices and let us define *mutual ergodicity* and *an ergodic index* as follows.

DEFINITION 1. For an $m \times m$ chain $\{A(k)\}$ of stochastic matrices, we say that an index $i \in [m]$ is ergodic if $\lim_{k \rightarrow \infty} A_i(k : t_0)$ exists for all $t_0 \geq 0$. Also, we say that two distinct indices $i, j \in [m]$ are mutually ergodic if $\lim_{k \rightarrow \infty} \|A_i(k : t_0) - A_j(k : t_0)\| = 0$.

From the definition it immediately follows that an index $i \in [m]$ is ergodic for a chain $\{A(k)\}$ if and only if $\lim_{k \rightarrow \infty} x_i(k)$ exists for any dynamics $\{x(k)\}$ driven by $\{A(k)\}$. Similarly, indices $i, j \in [m]$ are mutually ergodic if and only if $\lim_{k \rightarrow \infty} (x_i(k) - x_j(k)) = 0$ for any dynamics driven by $\{A(k)\}$.

The following result illustrates the relevance of the infinite flow graph to the study of the products of stochastic matrices.

LEMMA 1. ([8], Lemma 2) Two distinct indices $i, j \in [m]$ are mutually ergodic only if i and j belong to the same connected component of the infinite flow graph G^∞ of $\{A(k)\}$.

Generally, if i and j are mutually ergodic indices, it is not necessarily true that they are ergodic indices. As an example, consider the 4×4 stochastic chain $\{A(k)\}$ defined by:

$$A(2k) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad A(2k+1) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{for all } k \geq 1.$$

It can be verified that for any starting time $t_0 \geq 0$ and any $k > t_0$, we have $A(k : t_0) = A(k)$. Thus, it follows that indices 2 and 3 are mutually ergodic, while $\lim_{k \rightarrow \infty} A_2(k : t_0)$ and $\lim_{k \rightarrow \infty} A_3(k : t_0)$ do not exist.

The following result shows that under special circumstances, we can assert that some indices are ergodic if we know that a certain mutual ergodicity pattern exists in a chain.

LEMMA 2. Let S be a connected component of the infinite flow graph G^∞ of a chain $\{A(k)\}$. Suppose that indices i and j are mutually ergodic for all distinct $i, j \in S$. Then, every index $i \in S$ is ergodic.

PROOF. Without loss of generality let us assume that $S = \{1, \dots, i^*\}$ for some $i^* \in [m]$. Let \bar{S} be the complement of S . For the given chain $\{A(k)\}$ and the connected component S , let the chain $\{B(k)\}$ be defined by:

$$B_{ij}(k) = \begin{cases} A_{ij}(k) & \text{if } i \neq j \text{ and } i, j \in S \text{ or } i, j \in \bar{S}, \\ 0 & \text{if } i \neq j \text{ and } i \in S, j \in \bar{S} \text{ or } i \in \bar{S}, j \in S, \\ A_{ii}(k) + \sum_{\ell \in \bar{S}} A_{i\ell}(k) & \text{if } i = j \in S, \\ A_{ii}(k) + \sum_{\ell \in S} A_{i\ell}(k) & \text{if } i = j \in \bar{S}. \end{cases}$$

Then, $B(k)$ has the block diagonal structure of the following form

$$B(k) = \begin{bmatrix} B_1(k) & 0 \\ 0 & B_2(k) \end{bmatrix} \quad \text{for all } k \geq 1.$$

By construction the chain $\{B(k)\}$ is stochastic. It can be verified that $\sum_{k=1}^{\infty} |A_{ij}(k) - B_{ij}(k)| < \infty$ for all $i, j \in [m]$. Thus, $\{B(k)\}$ is an ℓ_1 -approximation of $\{A(k)\}$ as defined in [8]. Then, by Lemma 1 in [8], it follows that indices i and j are mutually ergodic for the chain $\{B(k)\}$ for all distinct $i, j \in S$. By the block diagonal form of $\{B(k)\}$, it follows that i and j are mutually ergodic for the $|S| \times |S|$

chain $\{B_1(k)\}$ and all $i, j \in S$. This, however, implies that the chain $\{B_1(k)\}$ is weakly ergodic (as defined in [3]) and, as proven in Theorem 1 in [3], this further implies that $\{B_1(k)\}$ is strongly ergodic, i.e. any index $i \in S$ is ergodic for $\{B_1(k)\}$. Again, by the application of Lemma 1 in [8], we conclude that any index $i \in S$ is ergodic for $\{A(k)\}$. **Q.E.D.**

2.2. Comparison Functions. Here, we show that under general conditions, a rich family of comparison functions exists for the dynamics $\{x(k)\}$ driven by a random chain $\{W(k)\}$.

Let us define an absolute probability process for an adapted chain $\{W(k)\}$, which is an extension of the concept of the absolute probability sequence introduced by A. Kolmogorov for deterministic chains in [4].

DEFINITION 2. *We say that a random (vector) process $\{\pi(k)\}$ is an absolute probability process for a random chain $\{W(k)\}$ adapted to $\{\mathcal{F}_k\}$ if*

1. *the random process $\{\pi(k)\}$ is adapted to $\{\mathcal{F}_k\}$,*
2. *the vector $\pi(k)$ is stochastic almost surely for all $k \geq 1$, and*
3. *the following relation holds almost surely*

$$\mathbb{E}[\pi^T(k+1)W(k+1) \mid \mathcal{F}_k] = \pi^T(k) \quad \text{for all } k \geq 0.$$

When an absolute probability process exists for a chain, we say that the chain admits an absolute probability process.

For a deterministic chain of stochastic matrices $\{A(k)\}$, Kolmogorov showed in [4] that there exists a sequence of stochastic vectors $\{v(k)\}$ such that $v^T(k+1)A(k+1) = v^T(k)$ for all $k \geq 0$. Note that, for an independent random chain, any absolute probability sequence for the expected chain is an absolute probability process for the random chain. Thus, the existence of an absolute probability process for an independent random chain of stochastic matrices follows immediately from the Kolmogorov's existence result. As another non-trivial example of random chains that admit an absolute probability process, one may consider an adapted random chain $\{W(k)\}$ that is doubly stochastic almost surely. In this case, the static sequence $\{\frac{1}{m}e\}$ is an absolute probability process for $\{W(k)\}$, where $e \in \mathbb{R}^m$ is the vector with all components equal to 1.

Now, suppose that we have an adapted chain $\{W(k)\}$ which admits an absolute probability sequence $\{\pi(k)\}$. Also, let $g : \mathbb{R} \rightarrow \mathbb{R}$ be an arbitrary convex function. Let us define the function $V_{g,\pi} : \mathbb{R}^m \times \mathbb{Z}^+ \rightarrow \mathbb{R}$, as follows:

$$(3) \quad V_{g,\pi}(x, k) = \sum_{i=1}^m \pi_i(k)g(x_i) - g(\pi^T(k)x) \quad \text{for all } x \in \mathbb{R}^m \text{ and all } k \geq 0.$$

From the definition of an absolute probability process, it follows that $V_{g,\pi}(x(k), k)$ is measurable with respect to \mathcal{F}_k for any dynamics $\{x(k)\}$ driven by a chain $\{W(k)\}$ that is adapted to $\{\mathcal{F}_k\}$. Also, since $\pi(k)$ is almost surely stochastic vector and g is a convex function, it follows that for any $x \in \mathbb{R}^m$, we have $V_{g,\pi}(x, k) \geq 0$ almost surely for all $k \geq 0$.

Next, we show that $V_{g,\pi}$ is a comparison function for the dynamics (2) for any convex function g . In particular, we prove that $\{V_{g,\pi}(x(k), k)\}$ is a super-martingale sequence irrespective of the initial point for the dynamics $\{x(k)\}$.

THEOREM 2. *Let $\{W(k)\}$ be an adapted chain that admits an absolute probability process $\{\pi(k)\}$. Then, for the dynamics (2) started at any initial point $(t_0, v) \in \mathbb{Z}^+ \times \mathbb{R}^m$, we have*

$$\mathbb{E}[V_{g,\pi}(x(k+1), k+1) \mid \mathcal{F}_k] \leq V_{g,\pi}(x(k), k) \quad \text{for all } k \geq t_0.$$

PROOF. By the definition of $V_{g,\pi}$ in (3), we have almost surely

$$\begin{aligned}
 V_{g,\pi}(x(k+1), k+1) &= \sum_{i=1}^m \pi_i(k+1)g(x_i(k+1)) - g(\pi^T(k+1)x(k+1)) \\
 &= \sum_{i=1}^m \pi_i(k+1)g([W(k+1)x(k)]_i) - g(\pi^T(k+1)x(k+1)) \\
 (4) \quad &\leq \sum_{i=1}^m \pi_i(k+1) \sum_{j=1}^m W_{ij}(k+1)g(x_j(k)) - g(\pi^T(k+1)x(k+1)),
 \end{aligned}$$

where in the second equality we use $[\cdot]_i$ to denote the i th component of a vector, while the inequality is obtained by using the convexity of $g(\cdot)$ and the fact that matrix $W(k)$ is stochastic almost surely. Since $\{\pi(k)\}$ is an absolute probability process for $\{W(k)\}$, it follows that $\mathbb{E}[\pi^T(k+1)W(k+1) \mid \mathcal{F}_k] = \pi^T(k)$. Also, since $x(k)$ is measurable with respect to \mathcal{F}_k , by taking the conditional expectation with respect to \mathcal{F}_k on both sides of Eq. (4), we obtain almost surely

$$\begin{aligned}
 \mathbb{E}[V_{g,\pi}(x(k+1), k+1) \mid \mathcal{F}_k] &\leq \sum_{j=1}^m \pi_j(k)g(x_j(k)) - \mathbb{E}[g(\pi^T(k+1)x(k+1)) \mid \mathcal{F}_k] \\
 &\leq \sum_{j=1}^m \pi_j(k)g(x_j(k)) - g(\mathbb{E}[\pi^T(k+1)x(k+1) \mid \mathcal{F}_k]),
 \end{aligned}$$

where the last inequality follows by the convexity of g and Jensen's inequality. The result follows by using $x(k+1) = W(k+1)x(k)$ and the definition of absolute probability process. **Q.E.D.**

Theorem 2 shows that the dynamics (2) admits infinitely many comparison functions, provided that $\{W(k)\}$ admits an absolute probability process.

Since $V_{g,\pi}(x(k), k) \geq 0$ almost surely for all $k \geq 0$, it follows that $\{V_{g,\pi}(x(k), k)\}$ is a bounded super-martingale. Hence, it is convergent almost surely irrespective of the initial point of the dynamics $\{x(k)\}$ and the choice of the convex function g .

COROLLARY 1. *Let $\{W(k)\}$ be an adapted chain that admits an absolute probability process $\{\pi(k)\}$. Then, for any dynamics $\{x(k)\}$ driven by $\{W(k)\}$ and for any convex function $g : \mathbb{R} \rightarrow \mathbb{R}$, the limit $\lim_{k \rightarrow \infty} V_{g,\pi}(x(k), k)$ exists almost surely.*

2.3. Quadratic Comparison Function. In the sequel, we focus on the particular choice of function $g(s) = s^2$ in relation (3). For convenience, we let

$$V_\pi(x, k) = \sum_{i=1}^m \pi_i(k)(x_i - \pi^T(k)x)^2 = \sum_{i=1}^m \pi_i(k)x_i^2 - (\pi^T(k)x)^2.$$

For this particular choice of the convex function g , we can provide a lower bound for the decrease of the conditional expectations $\mathbb{E}[V_g(x(k+1), k+1) \mid \mathcal{F}_k]$, which is exact under certain conditions.

THEOREM 3. *Let $\{W(k)\}$ be an adapted random chain with an absolute probability process $\{\pi(k)\}$. Then, for any dynamics $\{x(k)\}$ driven by $\{W(k)\}$, we have almost surely*

$$\mathbb{E}[V_\pi(x(k+1), k+1) \mid \mathcal{F}_k] \leq V_\pi(x(k), k) - \sum_{i < j} H_{ij}(k)(x_i(k) - x_j(k))^2 \quad \text{for all } k \geq t_0,$$

where $H(k) = \mathbb{E}[W^T(k+1)\text{diag}(\pi(k+1))W(k+1) \mid \mathcal{F}_k]$ with $\text{diag}(v)$ denoting the diagonal matrix induced by a vector v (i.e., with components v_i on the main diagonal), and $\sum_{i < j} = \sum_{i=1}^m \sum_{j=i+1}^m$. Furthermore, if $\pi^T(k+1)W(k+1) = \pi^T(k)$ almost surely, then the inequality holds as an equality.

PROOF. We have for all $k \geq t_0$,

$$(5) \quad V_\pi(x(k), k) = \sum_{i=1}^m \pi_i(k)x_i^2(k) - (\pi^T(k)x(k))^2 = x^T(k)\text{diag}(\pi(k))x(k) - (\pi^T(k)x(k))^2.$$

Thus, by letting $\Delta(x(k), k) = V_\pi(x(k), k) - V_\pi(x(k+1), k+1)$ and using $x(k+1) = W(k+1)x(k)$, we obtain for all $k \geq t_0$,

$$\begin{aligned} \Delta(x(k), k) &= x^T(k)\text{diag}(\pi(k))x(k) - (\pi^T(k)x(k))^2 \\ &\quad - \{x^T(k+1)\text{diag}(\pi(k+1))x(k+1) - (\pi^T(k+1)x(k+1))^2\} \\ &= x^T(k) [\text{diag}(\pi(k)) - W^T(k+1)\text{diag}(\pi(k+1))W(k+1)] x(k) \\ &\quad + \{(\pi^T(k+1)x(k+1))^2 - (\pi^T(k)x(k))^2\} \\ &= x^T(k)L(k)x(k) + \{(\pi^T(k+1)x(k+1))^2 - (\pi^T(k)x(k))^2\}, \end{aligned}$$

where $L(k) = \text{diag}(\pi(k)) - W^T(k+1)\text{diag}(\pi(k+1))W(k+1)$.

Note that the sequence $\{\pi^T(k)x(k)\}$ is a martingale, implying that $\{-(\pi^T(k)x(k))^2\}$ is a supermartingale. Thus, by taking the conditional expectation on both sides of the preceding equality and noticing that $x(k)$ is measurable with respect to \mathcal{F}_k , we have almost surely

$$(6) \quad \mathbb{E}[\Delta(x(k), k) \mid \mathcal{F}_k] \geq \mathbb{E}[x^T(k)L(k)x(k) \mid \mathcal{F}_k] = x^T(k)\mathbb{E}[L(k) \mid \mathcal{F}_k]x(k) \quad \text{for all } k \geq t_0.$$

Further, letting $e \in \mathbb{R}^m$ be the vector with all components equal to 1, from the definition of $L(k)$ we almost surely have for all $k \geq t_0$:

$$\begin{aligned} \mathbb{E}[L(k) \mid \mathcal{F}_k]e &= \mathbb{E}[\text{diag}(\pi(k))e - W^T(k+1)\text{diag}(\pi(k+1))W(k+1)e \mid \mathcal{F}_k] \\ &= \pi(k) - \mathbb{E}[W^T(k+1)\pi(k+1) \mid \mathcal{F}_k] = 0, \end{aligned}$$

which holds since $W(k)$ is stochastic almost surely and $\{\pi(k)\}$ is an absolute probability process for $\{W(k)\}$. Thus, the random matrix $\mathbb{E}[L(k) \mid \mathcal{F}_k]$ is symmetric and $\mathbb{E}[L(k) \mid \mathcal{F}_k]e = 0$ almost surely. It can be shown that for a symmetric matrix A with $Ae = 0$, we have $x^T Ax = -\sum_{i < j} A_{ij}(x_i - x_j)^2$. Then, it follows that almost surely

$$x^T(k)\mathbb{E}[L(k) \mid \mathcal{F}_k]x(k) = -\sum_{i < j} H_{ij}(k)(x_i(k) - x_j(k))^2,$$

where $H(k) = \mathbb{E}[W^T(k+1)\text{diag}(\pi(k+1))W(k+1) \mid \mathcal{F}_k]$. Using this relation in inequality (6), we conclude that almost surely

$$(7) \quad \mathbb{E}[V_\pi(x(k+1), k+1) \mid \mathcal{F}_k] \leq V_\pi(x(k), k) - \sum_{i < j} H_{ij}(k)(x_i(k) - x_j(k))^2 \quad \text{for all } k \geq t_0.$$

In the proof of inequality (7), the inequality appears due to relation (6) only. If $\pi^T(k+1)W(k+1) = \pi^T(k)W(k)$ almost surely, then we have $\pi^T(k+1)x(k+1) = \pi^T(k)x(k)$ almost surely. Thus, relation (6) holds as an equality, and consequently so does relation (7). **Q.E.D.**

One of the important implications of Theorem 3 is the following result.

COROLLARY 2. *Let $\{W(k)\}$ be an adapted random chain that admits an absolute probability process $\{\pi(k)\}$. Then, for any random dynamics $\{x(k)\}$ driven by $\{W(k)\}$, we have for all $t_0 \geq 0$,*

$$\mathbb{E} \left[\sum_{k=t_0}^{\infty} \sum_{i < j} L_{ij}(k) (x_i(k) - x_j(k))^2 \right] \leq \mathbb{E}[V_{\pi}(x(t_0), t_0)] < \infty,$$

where $L(k) = W^T(k+1) \text{diag}(\pi(k+1)) W(k+1)$.

PROOF. By taking expectation on both sides of the relation in Theorem 3, we obtain for all $k \geq t_0$:

$$(8) \quad \mathbb{E}[V_{\pi}(x(k+1), k+1)] \leq \mathbb{E}[V_{\pi}(x(k), k)] - \mathbb{E} \left[\sum_{i < j} \mathbb{E}[L_{ij}(k) | \mathcal{F}_k] (x_i(k) - x_j(k))^2 \right].$$

Since $x(k)$ is measurable with respect to \mathcal{F}_k , it follows that

$$\mathbb{E}[L_{ij}(k) | \mathcal{F}_k] (x_i(k) - x_j(k))^2 = \mathbb{E}[L_{ij}(k)(x_i(k) - x_j(k))^2 | \mathcal{F}_k] = \mathbb{E}[L_{ij}(k)(x_i(k) - x_j(k))^2].$$

Using this relation in Eq. (8), we see that for all $k \geq t_0$:

$$\mathbb{E}[V_{\pi}(x(k+1), k+1)] \leq \mathbb{E}[V_{\pi}(x(k), k)] - \mathbb{E} \left[\sum_{i < j} L_{ij}(k)(x_i(k) - x_j(k))^2 \right].$$

Hence, $\sum_{k=t_0}^{\infty} \mathbb{E} \left[\sum_{i < j} L_{ij}(k)(x_i(k) - x_j(k))^2 \right] \leq \mathbb{E}[V_{\pi}(x(t_0), t_0)]$ for any $t_0 \geq 0$. **Q.E.D.**

3. Class \mathcal{P}^* . In this section, we introduce a class of random chains which we refer to it as the class \mathcal{P}^* and prove one of the central results of this work. In particular, we show that the claim of Theorem 1 holds for any chain that is in the class \mathcal{P}^* and satisfies some form of aperiodicity.

DEFINITION 3. *The class \mathcal{P}^* is the class of random adapted chains that admit an absolute probability process $\{\pi(k)\}$ that is uniformly bounded away from zero almost surely, i.e., $\pi_i(k) \geq p^*$ almost surely for some scalar $p^* > 0$, and for all $k \geq 0$ and all $i \in [m]$. We write this concisely as $\{\pi(k)\} \geq p^* > 0$.*

It may appear that the definition of the class \mathcal{P}^* is rather a restrictive requirement. Later on, we show that in fact the class \mathcal{P}^* contains a broad family of deterministic and random chains.

To establish the main result of this section, we make use of the following intermediate result.

LEMMA 3. *Let $\{A(k)\}$ be a deterministic chain with the infinite flow graph $G^{\infty} = ([m], \mathcal{E}^{\infty})$. Let $(t_0, v) \in \mathbb{Z}^+ \times \mathbb{R}^m$ be an initial point for the dynamics driven by $\{A(k)\}$. If*

$$\lim_{k \rightarrow \infty} (x_{i_0}(k) - x_{j_0}(k)) \neq 0,$$

for some i_0, j_0 belonging to the same connected component of G^{∞} , then we have

$$\sum_{k=t_0}^{\infty} \sum_{i < j} [(A_{ij}(k+1) + A_{ji}(k+1))(x_i(k) - x_j(k))^2] = \infty.$$

PROOF. Let i_0, j_0 be in the same connected component of G^∞ with $\limsup_{k \rightarrow \infty} (x_{i_0}(k) - x_{j_0}(k)) = \alpha > 0$. Without loss of generality we may assume that $x(t_0) \in [-1, 1]^m$, for otherwise we can consider the dynamics started at $y(t_0) = \frac{1}{\|x(t_0)\|_\infty} x(t_0)$. Let S be the vertex set of the connected component in G^∞ containing i_0, j_0 , and without loss of generality assume that $S = \{1, 2, \dots, q\}$ for some $q \in [m]$, $q \geq 2$. Then, by the definition of the infinite flow graph, there exists a large enough $K \geq t_0$ such that

$$\sum_{k=K}^{\infty} A_S(k+1) \leq \frac{\alpha}{32q},$$

where $A_S(k+1) = A_{S\bar{S}}(k+1) + A_{\bar{S}S}(k+1)$. Furthermore, since $\limsup_{k \rightarrow \infty} (x_{i_0}(k) - x_{j_0}(k)) = \alpha > 0$, there exists a time instance $t_1 \geq K$ such that $x_{i_0}(t_1) - x_{j_0}(t_1) \geq \frac{\alpha}{2}$.

Let $\sigma : [q] \rightarrow [q]$ be a permutation such that $x_{\sigma(1)}(t_1) \geq x_{\sigma(2)}(t_1) \geq \dots \geq x_{\sigma(q)}(t_1)$, i.e. σ is an ordering of $\{x_i(t_1) \mid i \in [q]\}$. Since $x_{i_0}(t_1) - x_{j_0}(t_1) \geq \frac{\alpha}{2}$, it follows that $x_{\sigma(1)}(t_1) - x_{\sigma(q)}(t_1) \geq \frac{\alpha}{2}$ and, therefore, there exists $\ell \in [q]$ such that $x_{\sigma(\ell)}(t_1) - x_{\sigma(\ell+1)}(t_1) \geq \frac{\alpha}{2q}$. Let

$$T_1 = \arg \min_{t > t_1} \sum_{k=t_1}^t \sum_{\substack{i,j \in [q] \\ i \leq \ell, \ell+1 \leq j}} (A_{\sigma(i)\sigma(j)}(k+1) + A_{\sigma(j)\sigma(i)}(k+1)) \geq \frac{\alpha}{32q}.$$

Since S is a connected component of the infinite flow graph G^∞ , we must have $T_1 < \infty$; otherwise, S could be decomposed into two disconnected components $\{\sigma(1), \dots, \sigma(\ell)\}$ and $\{\sigma(\ell+1), \dots, \sigma(q)\}$.

Now, let $R = \{\sigma(1), \dots, \sigma(\ell)\}$. We have for any $k \in [t_1, T_1]$:

$$\begin{aligned} \sum_{k=t_1}^{T_1-1} A_R(k+1) &= \sum_{k=t_1}^{T_1-1} \left(\sum_{\substack{i,j \in [q] \\ i \leq \ell, \ell+1 \leq j}} (A_{\sigma(i)\sigma(j)}(k+1) + A_{\sigma(j)\sigma(i)}(k+1)) \right. \\ &\quad \left. + \sum_{i \leq \ell, j \in \bar{S}} A_{\sigma(i)j}(k+1) + \sum_{i \in \bar{S}, j \leq \ell} A_{i\sigma(j)}(k+1) \right) \\ &\leq \sum_{k=t_1}^{T_1-1} \sum_{\substack{i,j \in [q] \\ i \leq \ell, \ell+1 \leq j}} (A_{\sigma(i)\sigma(j)}(k+1) + A_{\sigma(j)\sigma(i)}(k+1)) + \sum_{k=K}^{\infty} A_S(k+1) \leq \frac{\alpha}{16q}, \end{aligned}$$

which follows by the definition of T_1 and the choice of $t_1 \geq K$. By Lemma 1 in [9], it follows that for $k \in [t_1, T_1]$,

$$\max_{i \in R} x_i(k) \leq \max_{i \in R} x_i(t_1) + 2\frac{\alpha}{16q}, \quad \min_{i \in S \setminus R} x_i(k) \geq \min_{i \in S \setminus R} x_i(t_1) - 2\frac{\alpha}{16q}.$$

Thus, for any $i, j \in [q]$ with $i \leq \ell$ and $j \geq \ell+1$, and for any $k \in [t_1, T_1]$, we have

$$x_{\sigma(i)}(k) - x_{\sigma(j)}(k) \geq 2 \left(2\frac{\alpha}{16q} \right) = \frac{\alpha}{4q}.$$

Therefore,

$$\begin{aligned} & \sum_{k=t_1}^{T_1} \sum_{\substack{i,j \in [q] \\ i \leq \ell, \ell+1 \leq j}} (A_{\sigma(i)\sigma(j)}(k+1) + A_{\sigma(j)\sigma(i)}(k+1))(x_{\sigma(i)}(k) - x_{\sigma(j)}(k))^2 \\ & \geq \left(\frac{\alpha}{4q}\right)^2 \sum_{k=t_0}^{T_1} \sum_{\substack{i,j \in [q] \\ i \leq \ell, j \geq \ell+1}} (A_{\sigma(i)\sigma(j)}(k+1) + A_{\sigma(j)\sigma(i)}(k+1)) \geq \left(\frac{\alpha}{4q}\right)^2 \frac{\alpha}{32q} = \beta > 0. \end{aligned}$$

Further, it follows that:

$$\begin{aligned} & \sum_{k=t_1}^{T_1} \sum_{i < j} (A_{ij}(k+1) + A_{ji}(k+1))(x_i(k) - x_j(k))^2 \\ & \geq \sum_{k=t_1}^{T_1} \sum_{\substack{i,j \in [q] \\ i \leq \ell, \ell+1 \leq j}} (A_{\sigma(i)\sigma(j)}(k+1) + A_{\sigma(j)\sigma(i)}(k+1))(x_{\sigma(i)}(k) - x_{\sigma(j)}(k))^2 \geq \beta. \end{aligned}$$

Since $\limsup_{k \rightarrow \infty} (x_{i_0}(k) - x_{j_0}(k)) = \alpha > 0$, there exists a time $t_2 > T_1$ such that $x_{i_0}(t_2) - x_{j_0}(t_2) \geq \frac{\alpha}{2}$. Then, using the above argument, there exists $T_2 > t_2$ such that $\sum_{k=t_2}^{T_2} \sum_{i < j} (A_{ij}(k+1) + A_{ji}(k+1))(x_i(k) - x_j(k))^2 \geq \beta$. Hence, using the induction, we can find time instances

$$\cdots > T_{\xi+1} > t_{\xi+1} > T_{\xi} > t_{\xi} > T_{\xi-1} > t_{\xi-1} > \cdots > T_1 > t_1 \geq t_0,$$

such that $\sum_{k=t_{\xi}}^{T_{\xi}} \sum_{i < j} (A_{ij}(k+1) + A_{ji}(k+1))(x_i(k) - x_j(k))^2 \geq \beta$ for any $\xi \geq 1$. The intervals $[t_{\xi}, T_{\xi}]$ are non-overlapping subintervals of $[t_0, \infty)$, implying that

$$\sum_{k=t_0}^{\infty} \sum_{i < j} (A_{ij}(k+1) + A_{ji}(k+1))(x_i(k) - x_j(k))^2 = \infty.$$

Q.E.D.

For our main result, let us define the weak aperiodicity for an adapted random chain.

DEFINITION 4. *We say that an adapted random chain $\{W(k)\}$ is weakly aperiodic if for some $\gamma > 0$, and for all distinct $i, j \in [m]$ and all $k \geq 0$,*

$$\mathbb{E}[W^{iT}(k+1)W^j(k+1) \mid \mathcal{F}_k] \geq \gamma \mathbb{E}[W_{ij}(k+1) + W_{ji}(k+1) \mid \mathcal{F}_k].$$

Now, we establish the main result of this section.

THEOREM 4. *Let $\{W(k)\} \in \mathcal{P}^*$ be an adapted chain that is weakly aperiodic. Then, $\lim_{k \rightarrow \infty} W(k : t_0) = W(\infty : t_0)$ exists almost surely for any $t_0 \geq 0$. Moreover, the event under which $W_i(\infty : t_0) = W_j(\infty : t_0)$ for all $t_0 \geq 0$ is almost surely equal to the event that i, j are belonging to the same connected component of the infinite flow graph of $\{W(k)\}$.*

PROOF. Since $\{W(k)\}$ is in \mathcal{P}^* , $\{W(k)\}$ admits an absolute probability process $\{\pi(k)\}$ such that $\{\pi(k)\} \geq p^* > 0$ almost surely. Thus, it follows that

$$p^* \mathbb{E}[W^T(k+1)W(k+1) \mid \mathcal{F}_k] \leq \mathbb{E}[W^T(k+1)\text{diag}(\pi(k+1))W(k+1) \mid \mathcal{F}_k] = H(k+1).$$

On the other hand, by the weak aperiodicity, we have

$$\gamma \mathbb{E}[W_{ij}(k+1) + W_{ji}(k+1) \mid \mathcal{F}_k] \leq \mathbb{E}[W^{iT}(k+1)W^j(k+1) \mid \mathcal{F}_k],$$

for some $\gamma \in (0, 1]$ and for all distinct $i, j \in [m]$. Thus, we have $p^* \gamma \mathbb{E}[W_{ij}(k+1) + W_{ji}(k+1) \mid \mathcal{F}_k] \leq H_{ij}(k+1)$. By Corollary 2, for the random dynamics $\{x(k)\}$ driven by $\{W(k)\}$ and started at arbitrary $(t_0, v) \in \mathbb{Z}^+ \times \mathbb{R}^m$, it follows that

$$p^* \gamma \sum_{k=t_0}^{\infty} \mathbb{E} \left[\sum_{i < j} (W_{ij}(k) + W_{ji}(k))(x_i(k) - x_j(k))^2 \right] \leq \mathbb{E}[V_{\pi}(x(t_0), t_0)].$$

As a consequence,

$$\sum_{k=t_0}^{\infty} \sum_{i < j} (W_{ij}(k) + W_{ji}(k))(x_i(k) - x_j(k))^2 < \infty \quad \text{almost surely.}$$

Therefore, by Lemma 3, we conclude that $\lim_{k \rightarrow \infty} (x_i(k, \omega) - x_j(k, \omega)) = 0$ for any i, j belonging to the same connected component of $G^{\infty}(\omega)$, for almost all $\omega \in \Omega$. By Lemma 2 it follows that every index $i \in [m]$ is ergodic for almost all $\omega \in \Omega$. By considering the initial conditions $(t_0, e_{\ell}) \in \mathbb{Z}^+ \times \mathbb{R}^m$ for all $\ell \in [m]$, the assertion follows. **Q.E.D.**

Theorem 4 shows that the dynamics in (2) is convergent almost surely for aperiodic chains $\{W(k)\} \in \mathcal{P}^*$. Moreover, the theorem also characterizes the limiting points of such a dynamics as well as the limit matrices of the products $W(k : t_0)$ as $k \rightarrow \infty$.

4. Balanced Chains. In this section, we characterize a subclass of \mathcal{P}^* chains, namely the class of *strongly aperiodic balanced chains*. We first show that this class includes many of the chains that have been studied in the existing literature. Then, we prove that any aperiodic balanced chain belongs to the class \mathcal{P}^* . We also show that a balanced independent random chain is strongly aperiodic, thus concluding Theorem 1.

Before continuing our analysis on balanced chains, let us discuss some of the well-known subclasses of such chains:

- 1. Balanced Bidirectional Chains:** We say that an independent chain $\{W(k)\}$ is a balanced bidirectional chain if there exists some $\alpha > 0$ such that $\mathbb{E}[W_{ij}(k)] \geq \alpha \mathbb{E}[W_{ji}(k)]$ for all $k \geq 1$ and $i, j \in [m]$. These chains are in fact balanced, since for any $S \subset [m]$ we have:

$$\mathbb{E}[W_{S\bar{S}}(k)] = \mathbb{E} \left[\sum_{i \in S, j \in \bar{S}} W_{ij}(k) \right] \geq \mathbb{E} \left[\sum_{i \in S, j \in \bar{S}} \alpha W_{ji}(k) \right] = \alpha \mathbb{E}[W_{\bar{S}S}(k)].$$

Examples of such chains are bounded bidirectional deterministic chains, which are the chains such that $A_{ij}(k) > 0$ implies $A_{ji}(k) > 0$ for all $i, j \in [m]$ and all $k \geq 1$, and the positive entries are uniformly bounded from below by some $\gamma > 0$ (i.e., $A_{ij}(k) > 0$ implies $A_{ij}(k) \geq \gamma$ for all $i, j \in [m]$ and all $k \geq 1$). In this case, for $A_{ij}(k) > 0$, we have $A_{ij}(k) \geq \gamma \geq \gamma A_{ji}(k)$ and for $A_{ij}(k) = 0$, we have $A_{ji}(k) = 0$ and, hence, in either of the cases $A_{ij}(k) \geq \gamma A_{ji}(k)$. Therefore, bounded bidirectional chains are examples of balanced bidirectional chains. Such chains have been considered in [6, 1, 2].

2. **Chains with Common Steady State** $\pi > 0$: This ensemble consists of independent random chains $\{W(k)\}$ such that $\mathbb{E}[\pi^T W(k)] = \mathbb{E}[\pi^T(k)]$ for some stochastic vector $\pi > 0$ and all $k \geq 1$, which are generalizations of doubly stochastic chains, where we have $\pi = \frac{1}{m}e$ (e is a vector of ones). Doubly stochastic chains and the chains with a common steady state $\pi > 0$ have been studied in [7, 9, 8].

To show that a chain with a common steady state $\pi > 0$ is a balanced chain, let us prove the following lemma.

LEMMA 4. *Let A be a stochastic matrix and $\pi > 0$ be a stochastic left-eigenvector of A corresponding to the unit eigenvalue, i.e., $\pi^T A = \pi^T$. Then, $A_{S\bar{S}} \geq \frac{\pi_{\min}}{\pi_{\max}} A_{\bar{S}S}$ for any non-trivial $S \subset [m]$, where $\pi_{\max} = \max_{i \in [m]} \pi_i$ and $\pi_{\min} = \min_{i \in [m]} \pi_i$.*

PROOF. Let $S \subset [m]$. Since $\pi^T A = \pi^T$, we have

$$(9) \quad \sum_{j \in S} \pi_j = \sum_{i \in [m], j \in S} \pi_i A_{ij} = \sum_{i \in S, j \in S} \pi_i A_{ij} + \sum_{i \in \bar{S}, j \in S} \pi_i A_{ij}.$$

On the other hand, since A is a stochastic matrix, we have $\pi_i \sum_{j \in [m]} A_{ij} = \pi_i$. Therefore,

$$(10) \quad \sum_{i \in S} \pi_i = \sum_{i \in S} \pi_i \sum_{j \in [m]} A_{ij} = \sum_{i \in S, j \in S} \pi_i A_{ij} + \sum_{i \in S, j \in \bar{S}} \pi_i A_{ij}.$$

Comparing Eq. (9) and Eq. (10), we see that $\sum_{i \in \bar{S}, j \in S} \pi_i A_{ij} = \sum_{i \in S, j \in \bar{S}} \pi_i A_{ij}$. Therefore,

$$\pi_{\min} A_{\bar{S}S} \leq \sum_{i \in \bar{S}, j \in S} \pi_i A_{ij} = \sum_{i \in S, j \in \bar{S}} \pi_i A_{ij} \leq \pi_{\max} A_{S\bar{S}}.$$

Hence, we have $A_{S\bar{S}} \geq \frac{\pi_{\min}}{\pi_{\max}} A_{\bar{S}S}$ for any non-trivial $S \subset [m]$. **Q.E.D.**

The above lemma shows that a chain with a common steady state $\pi > 0$ is balanced with balancedness coefficient $\alpha = \frac{\pi_{\min}}{\pi_{\max}}$. In fact, the lemma yields a much more general result, as provided below.

THEOREM 5. *Let $\{W(k)\}$ be an independent random chain with a sequence $\{\pi(k)\}$ of stochastic left-eigenvectors for the expected chain corresponding to the unit eigenvalue, i.e., $\pi^T(k) \mathbb{E}[W(k)] = \pi^T(k)$ for all $k \geq 1$. If $\{\pi(k)\} \geq p^*$ for some scalar $p^* > 0$, then $\{W(k)\}$ is a balanced chain with a balancedness coefficient $\alpha = \frac{p^*}{1-(m-1)p^*}$.*

PROOF. Since $\pi^T(k) \mathbb{E}[W(k)] = \pi^T(k)$ for all $k \geq 1$, by Lemma 4 we have

$$\mathbb{E}[W_{S\bar{S}}(k)] \geq \frac{\pi_{\min}(k)}{\pi_{\max}(k)} \mathbb{E}[W_{\bar{S}S}(k)] \quad \text{for any non-trivial } S \subset [m] \text{ and all } k \geq 1,$$

By $\{\pi(k)\} \geq p^* > 0$, it follows that $\pi_{\min}(k) \geq p^*$ for all $k \geq 1$. Since $\pi(k)$ is a stochastic vector, it further follows $\pi_{\max}(k) \leq 1 - (m-1)\pi_{\min}(k) \leq 1 - (m-1)p^*$. Therefore, for all $k \geq 1$,

$$\mathbb{E}[W_{S\bar{S}}(k)] \geq \frac{\pi_{\min}(k)}{\pi_{\max}(k)} \mathbb{E}[W_{\bar{S}S}(k)] \geq \frac{p^*}{1-(m-1)p^*} \mathbb{E}[W_{\bar{S}S}(k)],$$

for any non-trivial $S \subset [m]$. Thus, $\{W(k)\}$ is balanced with a balancedness coefficient $\alpha = \frac{p^*}{1-(m-1)p^*}$. **Q.E.D.**

Theorem 5 not only characterizes a class of balanced chains, but it also provides an alternative characterization of the balancedness for these chains. Thus, instead of verifying Definition 1 for every nontrivial subset $S \subset [m]$, for balancedness of independent random chains, it suffices to find a sequence $\{\pi(k)\}$ of stochastic (unit) left-eigenvectors of the expected chain $\{E[W(k)]\}$ such that the entries of the sequence do not vanish as time goes to infinity.

4.1. Absolute Probability Sequence for Balanced Chains. In this section, we show that any independent random chain that is strongly aperiodic and balanced must be in the class \mathcal{P}^* .

The road map to prove this result is as follows: we first show that this result holds for deterministic chains with uniformly bounded positive entries where by a uniformly bounded chain $\{A(k)\}$ we mean that there exists a scalar $\gamma > 0$ such that $A_{ij}(k) \geq \gamma$ whenever $A_{ij}(k) > 0$. Then, using this result and geometric properties of the set of strongly aperiodic balanced chains, we prove the statement for deterministic chains, which immediately implies the result for independent random chains.

To prove the result for uniformly bounded deterministic chains, we employ the technique that is used to prove Proposition 4 in [6]. However, the argument given in [6] needs some extensions to fit in our more general assumption of balancedness.

Let $\{A(k)\}$ be a deterministic chain of stochastic matrices. Let $S_j(k)$ be the set of indices corresponding to the positive entries in the j th column of $A(k : 0)$, i.e.,

$$S_j(k) = \{\ell \in [m] \mid A_{\ell j}(k : 0) > 0\} \quad \text{for all } j \in [m] \text{ and all } k \geq 0.$$

Also, let $\mu_j(k)$ be the minimum value of these positive entries, i.e.,

$$\mu_j(k) = \min_{\ell \in S_j(k)} A_{\ell j}(k : 0) > 0.$$

LEMMA 5. *Let $\{A(k)\}$ be a strongly aperiodic balanced chain such that the positive entries in each $A(k)$ are uniformly bounded from below by a scalar $\gamma > 0$. Then, $S_j(k) \subseteq S_j(k+1)$ and $\mu_j(k) \geq \gamma^{|S_j(k)|-1}$ for all $j \in [m]$ and $k \geq 0$.*

PROOF. Let $j \in [m]$ be arbitrary but fixed. By induction on k , we prove that $S_j(k) \subseteq S_j(k+1)$ for all $k \geq 0$ as well as the desired relation for $\mu_j(k)$. For $k = 0$, we have $A(0 : 0) = I$ by the definition, so $S_j(0) = \{j\}$. Then, $A(1 : 0) = A(1)$ and by the strongly aperiodic assumption on the chain $\{A(k)\}$ we have $A_{jj}(1) \geq \gamma$, implying $\{j\} = S_j(0) \subseteq S_j(1)$. Furthermore, we have $|S_j(0)| - 1 = 0$ and $\mu_j(0) = 1 = \gamma^0$. Hence, the claim is true for $k = 0$.

Now suppose that the claim is true for some $k \geq 0$, and consider $k+1$. Then, for any $i \in S_j(k)$, we have:

$$A_{ij}(k+1 : 0) = \sum_{\ell=1}^m A_{i\ell}(k+1)A_{\ell j}(k : 0) \geq A_{ii}(k+1)A_{ij}(k : 0) \geq \gamma\mu_j(k) > 0.$$

Thus, $i \in S_j(k+1)$, implying $S_j(k) \subseteq S_j(k+1)$.

To show the relation for $\mu_j(k+1)$, we consider two cases:

Case $A_{S_j(k)\bar{S}_j(k)}(k+1) = 0$: In this case for any $i \in S_j(k)$, we have:

$$(11) \quad A_{ij}(k+1 : 0) = \sum_{\ell \in S_j(k)} A_{i\ell}(k)A_{\ell j}(k : 0) \geq \mu_j(k) \sum_{\ell \in S_j(k)} A_{i\ell}(k+1) = \mu_j(k),$$

where the inequality follows from $i \in S_j(k)$ and $A_{S_j(k)\bar{S}_j(k)}(k+1) = 0$, and the definition of $\mu_j(k)$. Furthermore, by the balancedness of $A(k)$ and $A_{S_j(k)\bar{S}_j(k)}(k+1) = 0$, it follows that $0 =$

$A_{S_j(k)\bar{S}_j(k)}(k+1) \geq \alpha A_{\bar{S}_j(k)S_j(k)}(k+1) \geq 0$. Hence, $A_{\bar{S}_j(k)S_j(k)}(k+1) = 0$. Thus, for any $i \in \bar{S}_j(k)$, we have

$$A_{ij}(k+1:0) = \sum_{\ell=1}^m A_{i\ell}(k+1)A_{\ell j}(k:0) = \sum_{\ell \in \bar{S}_j(k)} A_{i\ell}(k+1)A_{\ell j}(k:0) = 0,$$

where the second equality follows from $A_{\ell j}(k:0) = 0$ for all $\ell \in \bar{S}_j(k)$. Therefore, in this case we have $S_j(k+1) = S_j(k)$, which by (11) implies $\mu_j(k+1) \geq \mu_j(k)$. In view of $S_j(k+1) = S_j(k)$ and the inductive hypothesis, we further obtain

$$\mu_j(k) \geq \gamma^{|S_j(k)|-1} = \gamma^{|S_j(k+1)|-1},$$

implying $\mu_j(k+1) \geq \gamma^{|S_j(k+1)|-1}$.

Case $A_{S_j(k)\bar{S}_j(k)}(k+1) > 0$: Since the chain is balanced, we have

$$A_{\bar{S}_j(k)S_j(k)}(k+1) \geq \alpha A_{S_j(k)\bar{S}_j(k)}(k+1) > 0,$$

implying that $A_{\bar{S}_j(k)S_j(k)}(k) > 0$. Therefore, by the uniform boundedness of $\{A(k)\}$, there exists $\hat{\xi} \in \bar{S}_j(k)$ and $\hat{\ell} \in S_j(k)$ such that $A_{\hat{\xi}\hat{\ell}}(k+1) \geq \gamma$. Hence, we have

$$A_{\hat{\xi}j}(k+1:0) \geq A_{\hat{\xi}\hat{\ell}}(k+1)A_{\hat{\ell}j}(k:0) \geq \gamma\mu_j(k) = \gamma^{|S_j(k)|},$$

where the equality follows by the induction hypothesis. Thus, $\hat{\xi} \in S_j(k+1)$ while $\hat{\xi} \notin S_j(k)$, which implies $|S_j(k+1)| \geq |S_j(k)| + 1$. This, together with $A_{\hat{\xi}j}(k+1:0) \geq \gamma^{|S_j(k)|}$, yields $\mu_j(k+1) \geq \gamma^{|S_j(k)|} \geq \gamma^{|S_j(k+1)|-1}$. **Q.E.D.**

The bound on $\mu_j(k)$ of Lemma 5 implies that the bound for the nonnegative entries given in Proposition 4 of [6] can be reduced from γ^{m^2-m+2} to γ^{m-1} .

Note that Lemma 5 holds for products $A(k:t_0)$ starting with any $t_0 \geq 0$, (with appropriately defined $S_j(k)$ and $\mu_j(k)$). An immediate corollary of Lemma 5 is the following result.

COROLLARY 3. *Under the assumptions of Lemma 5, we have for all $k > t_0 \geq 0$,*

$$\frac{1}{m}e^T A(k:t_0) \geq \min\left(\frac{1}{m}, \gamma^{m-1}\right)e^T,$$

where e is the vector of ones and the inequality is to be understood entry-wise.

PROOF. Without loss of generality, let us assume that $t_0 = 0$. Then, by Lemma 5 we have $\frac{1}{m}e^T A^j(k:0) \geq \frac{1}{m}|S_j(k)|\gamma^{|S_j(k)|-1}$ for any $j \in [m]$, where A^j denotes the j th column of A . For $\gamma \in [0, 1]$, the function $t \mapsto t\gamma^{t-1}$ defined on $[1, m]$ attains its minimum at either $t = 1$ or $t = m$. Therefore, $\frac{1}{m}e^T A(k:1) \geq \min(\frac{1}{m}, \gamma^{m-1})e^T$. **Q.E.D.**

Now, we relax the assumption on the bounded entries in Corollary 3.

THEOREM 6. *Let $\{A(k)\}$ be a balanced and strongly aperiodic chain. Then, there is a scalar $\gamma \in (0, 1]$ such that $\frac{1}{m}e^T A(k:0) \geq \min(\frac{1}{m}, \gamma^{m-1})e^T$ for all $k \geq 1$.*

PROOF. Let $\alpha > 0$ be a balancedness coefficient for $\{A(k)\}$ and let $A_{ii}(k) \geq \beta > 0$ for all $i \in [m]$ and $k \geq 1$. Further, let $\mathbf{B}_{\alpha,\beta}$ be the set of balanced matrices with the balancedness coefficient α and strongly aperiodic matrices with a coefficient $\beta > 0$, i.e.,

$$(12) \quad \mathbf{B}_{\alpha,\beta} := \{Q \in \mathbb{R}^{m \times m} \mid Q \geq 0, Qe = e, \\ Q_{S\bar{S}} \geq \alpha Q_{\bar{S}S} \text{ for all non-trivial } S \subset [m], Q_{ii} \geq \beta \text{ for all } i \in [m]\}.$$

The description in relation (12) shows that $\mathbf{B}_{\alpha,\beta}$ is a bounded polyhedral set in $\mathbb{R}^{m \times m}$. Let $\{Q^{(\xi)} \in \mathbf{B}_{\alpha,\beta} \mid \xi \in [n_{\alpha,\beta}]\}$ be the set of extreme points of this polyhedral set indexed by the positive integers between 1 and $n_{\alpha,\beta}$, which is the total number of extreme points of $\mathbf{B}_{\alpha,\beta}$.

Since $A(k) \in \mathbf{B}_{\alpha,\beta}$ for all $k \geq 1$, we can write $A(k)$ as a convex combination of the extreme points in $\mathbf{B}_{\alpha,\beta}$, i.e., there exist coefficients $\lambda_\xi(k) \in [0, 1]$ such that

$$(13) \quad A(k) = \sum_{\xi=1}^{n_{\alpha,\beta}} \lambda_\xi(k) Q^{(\xi)} \quad \text{with} \quad \sum_{\xi=1}^{n_{\alpha,\beta}} \lambda_\xi(k) = 1.$$

Now, consider the following independent random matrix process defined by:

$$W(k) = Q^{(\xi)} \quad \text{with probability } \lambda_\xi(k) \quad \text{for all } k \geq 1.$$

In view of this definition any sample path of $\{W(k)\}$ consists of extreme points of $\mathbf{B}_{\alpha,\beta}$. Thus, every sample path of $\{W(k)\}$ has a coefficient bounded by the minimum positive entry of the matrices in $\{Q^{(\xi)} \in \mathbf{B}_{\alpha,\beta} \mid \xi \in [n_{\alpha,\beta}]\}$, denoted by $\gamma = \gamma(\alpha, \beta) > 0$, where $\gamma > 0$ since $n_{\alpha,\beta}$ is finite. Therefore, by Corollary 3, we have $\frac{1}{m}e^T W(k : t_0) \geq \min(\frac{1}{m}, \gamma^{m-1})e^T$ for all $k > t_0 \geq 0$. Furthermore, by Eq. (13) we have $E[W(k)] = A(k)$ for all $k \geq 1$, implying

$$\frac{1}{m}e^T A(k : t_0) = \frac{1}{m}e^T E[W(k : t_0)] \geq \min\left(\frac{1}{m}, \gamma^{m-1}\right)e^T,$$

which follows from $\{W(k)\}$ being independent. **Q.E.D.**

Based on the above results, we are ready to prove the main result for deterministic chains.

THEOREM 7. *Any balanced and strongly aperiodic chain $\{A(k)\}$ is in the class \mathcal{P}^* .*

PROOF. As pointed out in [4] for any chain $\{A(k)\}$, there exists a sequence $\{t_r\}$ of time indices, such that for all $k \geq 0$, $\lim_{r \rightarrow \infty} A(t_r : k) = Q(k)$ exists and, for any stochastic vector $\pi \in \mathbb{R}^m$, the sequence $\{Q^T(k)\pi\}$ is an absolute probability sequence for $\{A(k)\}$. Since $\{A(k)\}$ is a balanced and strongly aperiodic chain, by Theorem 6 it follows that

$$\frac{1}{m}e^T Q(k) = \frac{1}{m} \lim_{r \rightarrow \infty} e^T A(t_r : k) \geq p^* e^T \quad \text{for all } k \geq 0,$$

with $p^* = \min(\frac{1}{m}, \gamma^{m-1}) > 0$. Thus, $\{\frac{1}{m}e^T Q(k)\}$ is a uniformly bounded absolute probability sequence for $\{A(k)\}$. **Q.E.D.**

The main result of this section follows immediately from Theorem 7.

THEOREM 8. *Any balanced and strongly aperiodic independent random chain is in the class \mathcal{P}^* .*

PROOF. The proof follows immediately by noticing that, for an independent random chain, $\{W(k)\}$, any absolute probability sequence for the expected chain $\{E[W(k)]\}$ is an absolute probability process for $\{W(k)\}$. **Q.E.D.**

As a result of Theorem 8 and Theorem 4, the proof of Theorem 1 follows immediately. In particular by Theorem 8, any independent random chain that is balanced and strongly aperiodic belongs to the class \mathcal{P}^* . Thus, the result follows by Theorem 4.

5. Connection to Non-negative Matrix Theory. In this section, we show that Theorem 1 is a generalization of the following well-known result in the non-negative matrix theory which plays a central role in the theory of ergodic Markov chains.

LEMMA 6. ([5], page 46) *For an aperiodic and irreducible stochastic matrix A , $\lim_{k \rightarrow \infty} A^k$ exists and it is equal to a rank one stochastic matrix.*

Recall that a stochastic matrix A is irreducible if there is no permutation matrix P such that

$$P^T A P = \begin{bmatrix} X & Y \\ \mathbf{0} & Z \end{bmatrix},$$

where X, Y, Z are $i \times i$, $i \times (m - i)$, and $(m - i) \times (m - i)$ matrices for some $i \in [m - 1]$ and $\mathbf{0}$ is the $(m - i) \times i$ matrix with all entries equal to zero.

Let us reformulate irreducibility using the tools we have developed in this paper.

LEMMA 7. *A stochastic matrix A is an irreducible matrix if and only if the static chain $\{A\}$ is balanced and its infinite flow graph is connected.*

PROOF. By the definition, a matrix A is irreducible if there is no permutation matrix P such that

$$P^T A P = \begin{bmatrix} X & Y \\ \mathbf{0} & Z \end{bmatrix}.$$

Since A is a non-negative matrix, we have that A is reducible if and only if there exists a subset $S = \{1, \dots, i\}$ for some $i \in [m - 1]$, such that

$$0 = [P^T A P]_{\bar{S}S} = \sum_{i \in \bar{S}, j \in S} e_i [P^T A P] e_j = \sum_{i \in \bar{S}, j \in S} A_{\sigma_i \sigma_j} = \sum_{i \in \bar{R}, j \in R} A_{ij},$$

where $\sigma_i = \{j \in [m] \mid P e_i = e_j\}$ (which is a singleton since P is a permutation matrix) and $R = \{\sigma_i \mid i \in S\}$. Thus, A is irreducible if and only if $A_{S\bar{S}} > 0$ for all non-trivial $S \subset [m]$. Therefore, by letting

$$\alpha = \min_{\substack{S \subset [m] \\ S \neq \emptyset}} \frac{A_{S\bar{S}}}{A_{\bar{S}S}},$$

and noting that $\alpha > 0$, we conclude that $\{A\}$ is balanced with a balancedness coefficient α . Furthermore, since $A_{S\bar{S}} + A_{\bar{S}S} \geq A_{S\bar{S}} > 0$ for all nontrivial $S \subset [m]$, it follows that the infinite flow graph of $\{A\}$ is connected.

Now, suppose that $\{A\}$ is balanced and its infinite flow graph of $\{A\}$ is connected. Then, $A_{S\bar{S}} > 0$ or $A_{\bar{S}S} > 0$ for all non-trivial $S \subset [m]$. By the balancedness of the chain it follows that $\min(A_{S\bar{S}}, A_{\bar{S}S}) > 0$ for any non-trivial $S \subset [m]$, implying that A is irreducible. **Q.E.D.**

Note that for an aperiodic A , we can always find some $h \geq 1$ such that $A_{ii}^h \geq \gamma > 0$ for all $i \in [m]$. Thus, based on Theorem 1, we have the following extension of Lemma 6 for independent random chains.

THEOREM 9. *Let $\{W(k)\}$ be a balanced and strongly aperiodic independent random chain with a connected infinite flow graph. Then, for any $t_0 \geq 0$, the product $W(k : t_0)$ converges to a rank one stochastic matrix almost surely (as k goes to infinity). Moreover, if $\{W(k)\}$ does not have the infinite flow property, the product $W(k : t_0)$ almost surely converges to a (random) matrix that has rank at most τ for any $t_0 \geq 0$, where τ is the number of connected components of the infinite flow graph of $\{E[W(k)]\}$.*

PROOF. The result follows immediately from Theorem 1. **Q.E.D.**

An immediate consequence of Theorem 9 is a generalization of Lemma 6 to inhomogeneous chains.

COROLLARY 4. *Let $\{A(k)\}$ be a balanced and strongly aperiodic stochastic chain. Then, $A(\infty : t_0) = \lim_{k \rightarrow \infty} A(k : t_0)$ exists for all $t_0 \geq 0$. Moreover, $A(\infty : t_0)$ is a rank one matrix for all $t_0 \geq 0$ if and only if the infinite flow graph of $\{A(k)\}$ is connected.*

REFERENCES

- [1] V.D. Blondel, J.M. Hendrickx, A. Olshevsky, and J.N. Tsitsiklis, *Convergence in multiagent coordination, consensus, and flocking*, Proceedings of IEEE CDC, 2005.
- [2] M. Cao, A. S. Morse, and B. D. O. Anderson, *Reaching a consensus in a dynamically changing environment: A graphical approach*, SIAM Journal on Control and Optimization **47** (2008), 575–600.
- [3] S. Chatterjee and E. Seneta, *Towards consensus: Some convergence theorems on repeated averaging*, Journal of Applied Probability **14** (1977), no. 1, 89–97.
- [4] A. Kolmogoroff, *Zur Theorie der Markoffschen Ketten*, Mathematische Annalen **112** (1936), no. 1, 155–160.
- [5] P.R. Kumar and P. Varaiya, *Stochastic systems estimation, identification and adaptive control*, Information and System Sciences Series, Prentice–Hall, Englewood Cliffs New Jersey, 1986.
- [6] J. Lorenz, *A stabilization theorem for continuous opinion dynamics*, Physica A: Statistical Mechanics and its Applications **355** (2005), 217223.
- [7] A. Nedić, A. Olshevsky, A. Ozdaglar, and J.N. Tsitsiklis, *On distributed averaging algorithms and quantization effects*, IEEE Transactions on Automatic Control **54** (2009), no. 11, 2506–2517.
- [8] B. Touri and A. Nedić, *On approximations and ergodicity classes in random chains*, Under review, 2010.
- [9] B. Touri and A. Nedić, *On ergodicity, infinite flow and consensus in random models*, IEEE Transactions on Automatic Control **56** (2011), no. 7, 1593–1605.

DEPT. OF INDUSTRIAL AND ENTERPRISE SYSTEMS ENGINEERING, 104 S. MATHEWS AVE., URBANA, IL 61801, USA